Conjoint Approach of the "Residual" Prediction and the Nonlinear Autoregressive Neural Network Increases the Forecast Precision of the Base Model

Alexander Sergeev[1,2, a)], Andrey Shichkin[1,2, b)], Alexander Buevich[1,2, c)], Elena Baglaeva[1,2, d)], Irina Subbotina[2, e)], Anna Rakhmatova[1,2, f)], Alexandra Kosachenko[1, g)], Anastasia Moskaleva[1, h)], Alexander Medvedev[1,2, i)], and Marina Sergeeva[2, j)]

[1]Ural Federal University, Mira str., 19, Ekaterinburg, RUSSIA 620002
[2]Institute of Industrial Ecology UB RAS, S. Kovalevskoy str., 20, Ekaterinburg, RUSSIA 620990

[a)]aleksandrpsergeev@gmail.com
[b)]and@ecko.uran.ru
[c)]bagalex3@gmail.com
[d)]elenbaglaeva@gmail.com
[e)]iesub@mail.ru
[f)]anya_rahmatova@mail.ru
[g)]alleshch7@gmail.com
[h)]amoskalyova03@gmail.com
[i)]Corresponding author: medvedev@ecko.uran.ru
[j)]marin@ecko.uran.ru

Abstract. An algorithm based on predicting the residuals of the nonlinear autoregressive neural network model with external input (NARX), which can improve the prediction accuracy, was proposed. Data of the concentration of one of the main greenhouse gases methane (CH_4) on the Arctic Island of Belyy, Russia, were used for prediction. A time interval, which was characterized by high daily fluctuations in the CH_4 concentration was selected. The forecast accuracy was determined by the mean absolute error (MAE), root mean squared error (RMSE) and root mean squared relative error (RMSRE) errors. The use of the algorithm allowed to increase the forecast accuracy from 11% for RMSE to 20% for RMSRE.

INTRODUCTION

Forecasts of the time series are very in demand in the studies concerning the problem of the global warming. This is especially noticeable in the arctic areas. For example, the average temperature in some areas of the Russian part of the Arctic was 6 degrees higher than the average long-term values. [1]-[2]. Forecasting the dynamics of atmospheric air pollution is carried out both with using of classical statistical approaches [3]-[8] and models based on artificial neural networks, which have become particularly popular in recent years [9]-[16]. Among the many types of artificial neural networks with time series prediction problems the Nonlinear Autoregressive Neural Network with an External Input (NARX) network is the most suitable [17]-[20].

NARX network is a recurrent dynamic network consisting of several layers and having a feedback. NARX network is based on the autoregressive model used to describe systems that have inertia. The value of the indicator (ppm mole fraction CH_4), corresponding to the time, applied to the input of the network, as the output, were used.

The goal of our work was to compare the most used types of ANN for forecasting the time series.

International Conference of Numerical Analysis and Applied Mathematics ICNAAM 2019
AIP Conf. Proc. 2293, 120021-1–120021-4; https://doi.org/10.1063/5.0027179
Published by AIP Publishing. 978-0-7354-4025-8/$30.00

MATERIALS AND METHODS

Measurements of greenhouse gases methane, carbon dioxide, carbon monoxide, and water vapor were made on the Arctic Island Belyy, YNAO, Russia in summer 2015. The island is located in the Kara Sea 5-10 km north of the Yamal Peninsula (Fig. 1).

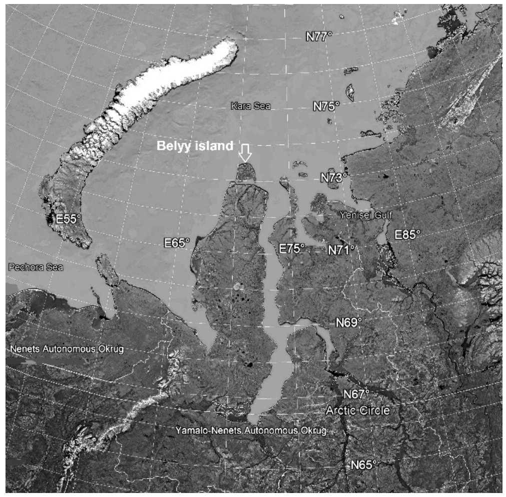

FIGURE 1. Place of measurements (Google Earth).

Concentration of greenhouse gases were measured by means of a cavity ring-down spectrometer Picarro G2401. The main meteorological parameters (temperature, humidity, atmospheric pressure) were measured by the Vaisala AWS310 weather station. Data of the gas concentrations and meteorological parameters were synchronized. In this work, time series of 192 intervals (hours) were chosen. For the prediction data of the methane content (ppm mole fraction) in the surface layer of atmospheric air were used. For modeling, the data was divided into two subsets. The first was used to train the network). Traditionally, this subset includes most of the test sample (168 intervals in our case). The remaining 24 intervals formed a test subset. It was not used in the learning process of the neural network; and was used only to predict the values of the series.

The networks structure was determined during computer simulation. Values of the ppm mole fraction CH_4, basic meteorological parameters (temperature, humidity, atmospheric pressure) were the input data. The hidden layer consisted of a several neurons, and the output layer represented the element concentration (ppm mole fraction CH_4) corresponding to the time. For ANNs the Levenberg-Marquardt training algorithm was used. The selection of the number of neurons in the hidden layer in NARX was carried out by the lower root mean squared error (RMSE) (4). The number of neurons was varied from 5 to 25. Each network was trained 500 times and the best of them was selected.

The hybrid approach was realized by the three-step algorithm combining two same interpolation techniques in one ensemble. Trained NARX network predicted content in the training data set points. Then the residuals at the same points were computed. Residuals in the neural network were defined as follows:

$$r(\boldsymbol{t}_i) = p_{ANN}(\boldsymbol{t}_i) - o(\boldsymbol{t}_i), \tag{1}$$

where $r(t_i)$ are the residuals of the data set ti, $o(\boldsymbol{t}_i)$ are the measured values, $p_{ANN}(\boldsymbol{t}_i)$ are the values estimated by the neural network. The resulting residuals are estimated using NARX network. The NARX network was applied in order to predict the residuals. These residuals were input data for NARX network. The same learned ANNs calculated element content values in the test data set points.

The final step was estimation produced as a sum of NARX network predictions and NARX network estimates of the residuals. The element content $C_{CH4}(\mathbf{t}_i)$ was obtained as the sum of NARX network evaluation and residuals evaluation by NARX network:

$$C_{CH4}(\mathbf{t}_i) = p_{ANN}(\boldsymbol{t}_i) + r(\boldsymbol{t}_i), \tag{2}$$

The indices mean absolute error MAE (3), RMSE (4), and a root mean squared relative error (RMSRE) (5) was verified the predictive accuracy of each selected approach between the prediction and raw data from the training data set.

$$MAE = \frac{\sum_{i=1}^{n} |z_{mod}(x_i) - z(x_i)|}{n}, \tag{3}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(z_{mod}(x_i)-z(x_i))^2}{n}}, \quad (4)$$

$$RMSRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{z_{mod}(x_i)-z(x_i)}{x_i}\right)^2}, \quad (5)$$

where $z_{mod}(\boldsymbol{x}_i)$ is a predicted concentration in location $\boldsymbol{x}_i$, $z(\boldsymbol{x}_i)$ is a measured concentration, and n is a number of points.

RESULTS AND DISCUSSION

The final neuron number in the hidden layer was 20 for NARX network. Table 1 shows the parameters used to compare the performance of the different methods (the best values demonstrated by NARXR for the test interval are in **bold**).

TABLE 1. Accuracy assessment indices of the CH4 concentration.

Subset	Method	MAE, ppm	RMSE, ppm	RMSRE, ppm
Training	NARX	0.005	0.007	0.004
Training	NARXR	0.006	0.007	0.004
Test	NARX	0.007	0.009	0.005
Test	NARXR	**0.006**	**0.008**	**0.004**

Analysis of the residuals showed that the model did not fully identify the likely patterns that may be present in the data. The correlation coefficient between the original data and the residuals was 0.6. In this case, it is advisable to apply a hybrid technique and re-model the residuals by NARX. Figure 2 shows the forecast results for two models.

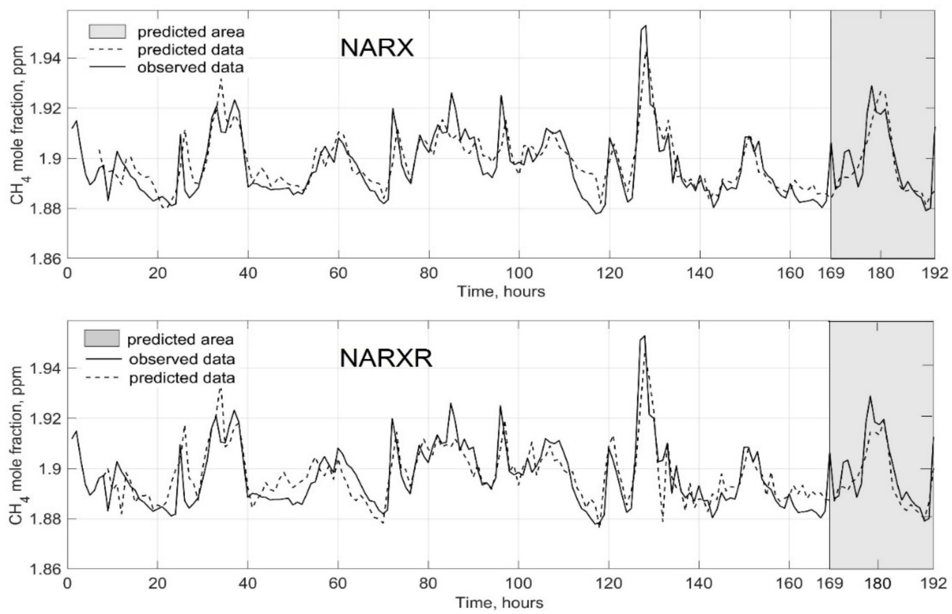

FIGURE 2. Comparison of different prediction approaches

As a result of applying the hybrid approach, it was possible to improve the prediction accuracy for all indices. So, for MAE, the improvement was 15%, for RMSE by 11%. For RMSRE, the improvement was the most significant and amounted to 20%. At the same time for a training set the models accuracy were comparable.

CONCLUSION

The paper presents a comparison of the models based on NARX networks for the time series prediction on the data obtained on the Arctic island Belyy. For the study, concentrations of one of the main greenhouse gases, methane, were taken. The time interval of 168 hours was used for ANN training; the forecast followed for the next 24-hour interval. A new approach that improved the prediction accuracy is proposed. A hybrid approach based on NARX and analysis of its prediction residues has demonstrated its effectiveness for our data. The algorithm is quite simple and can be used to improve the prediction of time series in different studies.

REFERENCES

[1] C. Nullis, Provisional WMO Statement on the Status of the Global Climate in 2016, World Meteorological Organization (2016).

[2] M.C. Serreze and R.G. Barry, "Processes and impacts of Arctic amplification: A research synthesis", Global and Planetary Change 77, 85–96 (2011).

[3] N. Chu, J.B. Kadane, C.I. Davidson, "Using statistical regressions to identify factors influencing PM2.5 concentrations: the Pittsburgh supersite as a case study", Aerosol Sci. Technol. 44, 766-774 (2010).

[4] W.G. Cobourn, "An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations", Atmos. Environ. 44, 3015-3023 (2010).

[5] A. Erdil and E. Arcaklioglu, "The prediction of meteorological variables using artificial neural network", Neural Computing & Application 22, 1677–1683 (2013).

[6] Genc, D.D., Yesilyurt, C., Tuncel, G., "Air pollution forecasting in Ankara, Turkey using air pollution index and its relation to assimilative capacity of the atmosphere", Environ. Monit. Assess. 166, 11-27 (2010).

[7] A. Russo, F. Raischel and P. Lind, "Air quality prediction using optimal neural networks with stochastic variables", Atmos. Environ. 79, 822–830 (2013).

[8] Q. Zhou, H. Jiang, J. Wang and J. Zhou, "A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network", Sci. Total Environ. 496, 264–274 (2014).

[9] G. Panayotova, G. Dimitrov, P. Petrov and O. Bychkov, "Modeling and data processing of information systems", in 3rd International Conference on Artificial Intelligence and Pattern Recognition, (Lodz, Poland, IEEE, 2016), pp. 154-158.

[10] A.P. Sergeev, A.G. Buevich, E.M. Baglaeva and A.V. Shichkin, "Combining spatial autocorrelation with machine learning increases prediction accuracy of soil heavy metals", Catena 174, 425–435 (2019).

[11] D.A. Tarasov, A.G. Buevich, A.P. Sergeev and A.V. Shichkin, "High Variation Topsoil Pollution Forecasting in the Russian Subarctic: Using Artificial Neural Networks Combined with Residual Kriging", Applied Geochemistry 88(B), 188-197, (2018).

[12] J. Vasilev, V.F. Turygina, A.I. Kosarev and Y.Y. Nazarova, "Mathematical optimization in environmental economics. Algorithm of gradient projection method, in International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM, 2016, pp. 349-356.

[13] G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model", Neurocomputing 5, 159–175 (2003).

[14] Andrey Shichkin, Alexander Buevich, Alexander Sergeev, Alexander Medvedev, Alexandra Kosachenko, Anastasia Moskaleva, Marina Sergeeva. Training Algorithms for Artificial Neural Networks for Time Series Forecasting of Greenhouse Gas Concentrations. AIP Conference Proceedings 2116, 200019 (2019).

[15] Sergeev A.P., Buevich A.G., Medvedev A.N., Spasov K.B., Kosachenko A.I., Moskaleva A.S. On the Hybrid Models of Soil Contaminants Concentrations Predicting in Subarctic Region. AIP Conference Proceedings 1978, 440006 (2018).

[16] D.A. Tarasov, A.N. Medvedev, A.P. Sergeev, and A.G. Buevich, Review and Possible Development Direction of the Methods for Modeling of Soil Pollutants Spatial Distribution. AIP Conference Proceedings 1863, 050014 (2017).

[17] M. Ardalani-Farsa and S. Zolfaghari, "Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks", Neurocomputing 73, 2540–2553 (2010).

[18] J.M. Jr. Menezes and G.A. Barreto, "Long-term time series prediction with the NARX network: An empirical evaluation", Neurocomputing 71, 3335–3343 (2008).

[19] E. Pisoni, M. Farina, C. Carnevale and L. Piroddi, "Forecasting peak air pollution levels using NARX models", Engineering Applications of Artificial Intelligence 22, 593–602 (2009).

[20] R. Zemouri, R. Gouriveau, N. Zerhouni, "Defining and applying prediction performance metrics on a recurrent NARX time series model", Neurocomputing 73, 2506–2521 (2010).